

# Cosmic Shear as a Probe of Galaxy Formation Physics

Simon Foreman<sup>★1,2,3,4</sup>, Matthew R. Becker<sup>1,2</sup>, Risa H. Wechsler<sup>1,2,4</sup>

<sup>1</sup>*Kavli Institute for Particle Astrophysics and Cosmology, P.O. Box 2450, Stanford, CA 94305*

<sup>2</sup>*SLAC National Accelerator Laboratory, 2575 Sand Hill Road, Menlo Park, CA 94025*

<sup>3</sup>*Stanford Institute for Theoretical Physics, 382 Via Pueblo Mall, Stanford, CA 94305*

<sup>4</sup>*Department of Physics, Stanford University, 382 Via Pueblo Mall, Stanford, CA 94305*

31 May 2016

## ABSTRACT

We evaluate the potential for current and future cosmic shear measurements from large galaxy surveys to constrain the impact of baryonic physics on the matter power spectrum. We do so using a model-independent parameterization that describes deviations of the matter power spectrum from the dark-matter-only case as a set of principal components that are localized in wavenumber and redshift. We perform forecasts for a variety of current and future datasets, and find that at least  $\sim 90\%$  of the constraining power of these datasets is contained in no more than nine principal components. The constraining power of different surveys can be quantified using a figure of merit defined relative to currently available surveys. With this metric, we find that the final Dark Energy Survey dataset (DES Y5) and the Hyper Suprime Cam Survey will be roughly an order of magnitude more powerful than existing data in constraining baryonic effects. Upcoming Stage IV surveys (LSST, Euclid, and WFIRST) will improve upon this by a further factor of a few. We show that this conclusion is robust to marginalization over several key systematics. The ultimate power of cosmic shear to constrain galaxy formation is dependent on understanding systematics in the shear measurements at small (sub-arcminute) scales. If these systematics can be sufficiently controlled, cosmic shear measurements from DES Y5 and other future surveys have the potential to provide a very clean probe of galaxy formation and to strongly constrain a wide range of predictions from modern hydrodynamical simulations.

**Key words:** gravitational lensing; weak – cosmology; observations – galaxies: formation

## 1 INTRODUCTION

Weak gravitational lensing has tremendous potential to inform our knowledge of the universe, on distance scales ranging from galactic to cosmological. In particular, *cosmic shear*, the statistical measurement of the distortion of observed galaxy shapes due to lensing by the distribution of matter along the line of sight, can provide information about the background cosmological model, via its sensitivity to the growth of structure and the strength of gravitational clustering on various scales. The first successful measurements of cosmic shear were carried out using patches of the sky with area of order one square degree (Bacon et al. 2000; Kaiser et al. 2000; Wittman et al. 2000; Van Waerbeke et al. 2000). More recent measurements from the Stripe 82 region of the Sloan Digital Sky Survey (Lin et al. 2012; Huff et al. 2014), the Canada France Hawaii Telescope Lensing Survey (e.g. Heymans et al. 2013), the Dark Energy Survey Sci-

ence Verification run (The Dark Energy Survey Collaboration et al. 2015), and the Kilo-Degree Survey (Kuijken et al. 2015) have used  $\mathcal{O}(100)$  square degrees and are beginning to provide useful low-redshift cosmological constraints that are complementary to the high-redshift information obtained from the cosmic microwave background. Indeed, there currently exist discrepancies between the two sets of constraints that have yet to be resolved (see e.g. MacCrann et al. 2015 for further discussion). Future surveys like the Wide Field InfraRed Survey Telescope (WFIRST)<sup>1</sup>, the Large Synoptic Survey Telescope (LSST)<sup>2</sup>, and Euclid<sup>3</sup> will expand the sky coverage to  $\mathcal{O}(10000)$  square degrees, and measure the shapes of orders of magnitude more galaxies than previous observational programs.

The observed shapes of galaxies are altered by gravitational lensing by only about 1%, and it is the statistical

★ E-mail: sfore@stanford.edu

<sup>1</sup> <http://wfIRST.gsfc.nasa.gov>

<sup>2</sup> <http://www.lsST.org>

<sup>3</sup> <http://sci.esa.int/euclid>

correlations among these shapes that are the primary end product of cosmic shear observations. There are many possible sources of systematic errors that could contaminate the underlying cosmological signal, and each must be characterized in detail in order to obtain robust results. These include instrumental distortions of observed galaxy shapes caused by the telescope’s point-spread function; biases in the methods used to measure shapes from raw images; uncertainties in the photometric redshifts obtained for each galaxy; “intrinsic alignments” of galaxies with nearby structures, independently of the alignments caused by lensing by the line-of-sight distribution of matter; and uncertainties in the theoretical modeling of the distribution of matter on small scales.

This paper will focus on the last point, in particular on the effects of baryonic physics on the underlying two-point statistics of the distribution of matter. It is well-known (White 2004; Zhan & Knox 2004; Jing et al. 2006; Rudd et al. 2008) that these effects can strongly influence cosmological constraints derived from cosmic shear measurements if not properly accounted for. For this reason, it is common to in cosmological analyses to simply exclude scales where these effects are important. Alternatively, a variety of schemes have been developed to allow inclusion of these scales by mitigating the effects of baryons on the desired information. These includes parametrizations of baryonic effects in the context of the halo model (Zentner et al. 2008; Semboloni et al. 2011, 2013; Zentner et al. 2013; Mohammed et al. 2014; Mohammed & Seljak 2014; Mead et al. 2015), perturbation theory (Lewandowski et al. 2015), principal component analysis (Eifler et al. 2015; Kitching et al. 2016), or empirical fitting functions for the net effect (Harnois-Déraps et al. 2015) or individual baryonic components (Schneider & Teyssier 2015), along with other techniques not based on direct parametrizations (Huterer & White 2005; Kitching & Taylor 2011; Bielefeld et al. 2015).

However, one can also turn the question around, and treat these baryonic effects as *signal* to be constrained instead of nuisance to be removed or marginalized over. The advantage of using cosmic shear for these constraints is that it provides a relatively clean probe of the underlying matter distribution, free of the need to classify individual objects and explicitly relate them to dark matter halos or properties of their environments. Similarly, the effect of baryons on the total matter distribution can also be obtained relatively easily from a numerical simulation, by simply measuring the  $N$ -point matter statistics from snapshots of the simulation output. In recent years, such simulations have reached sufficient levels of detail that it is worthwhile to compare their predictions to observations of the real universe, and cosmic shear can act as nice complement to other observations that have already been exploited for this purpose (e.g. Schaye et al. 2010; Viola et al. 2015).

Rather than performing a separate analysis to constrain each implementation of baryonic physics one is interested in, it is possible to obtain generic, model-independent constraints via a principal component analysis of the deviations of the observed matter power spectrum from a fiducial “dark matter-only” case. One can construct principal components of these deviations (in the form of linear combinations of deviations from the dark matter-only power spectrum at discrete points in the wavenumber-redshift plane) from a Fisher

matrix corresponding to a specific survey: inverting and then diagonalizing this matrix not only gives the principal components corresponding to that survey, but also provides an automatic ranking of these principal components in terms of the expected constraints on their amplitudes. (For applications of this procedure to other problems in cosmology, see Huterer & Starkman 2003; Albrecht et al. 2009; Samsing et al. 2012.) One can then simply constrain these amplitudes in a likelihood analysis, and thereafter straightforwardly relate these constraints to any baryonic model of interest.

One may contrast this approach to that of Eifler et al. (2015), which also applies principal component analysis to the issue of baryonic effects. The approach in that work is to find the principal components of the variations in the matter power spectrum between a range of specific baryonic models, and then remove these principal components from both the data and theory vectors in a likelihood analysis, in order to minimize the amount that cosmological constraints are contaminated by these effects. Our approach is rather to find the principal components of the matter power spectrum that will be best constrained by a cosmic shear measurements from a given survey, and then relate constraints on these principal components to constraints on a set of baryonic models as a second independent step.

In this paper, we carry out forecasts to assess the performance of this approach when it is applied to current and upcoming weak lensing surveys. Our forecasts marginalize over a variety of systematic effects relating both to calibration of the shear measurements and uncertainties in the theoretical modeling. For all surveys we consider, we find that 90% of the constraining power contained in the shear correlation function is captured by at most nine principal components of the matter power spectrum. We use this fact to define a figure of merit: the reciprocal of the geometric mean on the expected one-sigma constraints of the amplitudes of the nine principal components. Comparing this figure of merit for different surveys, we find that the constraints on baryonic effects from Stage III surveys like the Dark Energy Survey or Hyper Suprime Cam Survey<sup>4</sup> can improve upon those from currently available datasets by roughly an order of magnitude, while Stage IV surveys will only improve upon this by a factor of a few, with the constraining power being driven mainly by the number of galaxy shapes that can be measured by each survey.

We then apply this method to a representative set of models for baryonic effects on the matter power spectrum; specifically, we use nine models from the Overwhelmingly Large Simulations (OWLS; Schaye et al. 2010; Van Daalen et al. 2011) suite. We find that the constraining power of future surveys is strongly dependent on the minimum angular scale at which measurements of the shear correlation functions can be used. In particular, measurements at scales of less than one arcminute would allow Stage III surveys to rule out the majority of the OWLS models at more than five-sigma confidence (or alternatively rule out the dark matter-only power spectrum if one of these models were correct). Furthermore, these measurements would also likely allow for the identification of a single “best-fit” model among those that currently exist.

<sup>4</sup> <http://www.naoj.org/Projects/HSC/>

While we marginalize over a wide range of systematics, our conclusions rely on the existence of shear catalogs with well-understood properties (in terms of, for example, selection functions) at the appropriate scales. Provided that such an understanding can be attained (particularly for  $\theta \lesssim 1'$ ), however, our results demonstrate that cosmic shear is capable of providing stringent tests of the results of current and future hydrodynamical simulations, opening an additional window on the physics of galaxy formation and evolution (or at least, its implementation in simulations).

The paper is organized as follows. In Sec. 2, we describe the principal component method we use for parameterizing the matter power spectrum, while in Sec. 3 we describe our assumptions about the surveys we provide forecasts for, the mock cosmic shear data and covariances we associate with these surveys, and the set of baryonic models we will use for our tests. In Sec. 4, we verify that most of the constraining power of each survey with respect to the matter power spectrum is contained in no more than the first nine principal components, and describe how we relate constraints on these principal components to a given baryonic model. In Sec. 5, we compare the performance of various surveys via an appropriately-defined figure of merit, and also quantify this performance in terms of constraints on the baryonic models in our test set. We additionally examine the impact of varying our priors on the most important theoretical systematics. Finally, we provide additional comments and conclude in Sec. 6.

## 2 METHOD OF PRINCIPAL COMPONENTS

We begin by seeking a generalized parametrization of the matter power spectrum that allows for deviations from the DM-only case that are localized in wavenumber  $k$  and redshift  $z$ . Such a parametrization will allow us to account for the fact that the strongest constraints on  $P(k, z)$  from cosmic shear will be localized where the lensing kernels peak, and also for the fact that models for baryonic effects will have characteristic scale- and redshift-dependence that will vary from model to model.

One such parametrization simply uses the fractional deviation  $\Delta$  between the full power spectrum and the DM-only one at discrete values of  $k$  and  $z$ , taking the value of that deviation at each sample as a free parameter and smoothly interpolating between these samples at other values of  $k$  and  $z$ :

$$P(k_i, z_j) = P_{\text{DM-only}}(k_i, z_j) [1 + \Delta_{ij}] . \quad (1)$$

A version of this parametrization that assumes that  $\Delta$  is  $z$ -independent was used in Huterer & Takada (2005) and Hearin et al. (2012). (For other works that do not rely on a  $z$ -independent parametrization, see Bernstein 2009; Bielefeld et al. 2015.) For our purposes, it is important to relax this assumption, but this requires the introduction of at least several tens of new parameters  $\{\Delta_{ij}\}$  into any analysis that would make use of this parametrization. Following the discussion above, some of these parameters will be much better constrained by a particular dataset than others. In addition, the nature of cosmic shear observables as projections of the matter power spectrum will lead to strong degeneracies among these parameters, since there are many ways to

perturb the  $\{\Delta_{ij}\}$  in a correlated way without significantly affecting the final signal.

Therefore, we further seek a convenient basis for these parameters that eliminates degeneracies as much as possible, and also identifies which basis elements are most relevant for a given set of observations. In fact, such a basis is straightforward to construct: given a covariance matrix for the  $\{\Delta_{ij}\}$ , its eigenvectors would specify a set of linear combinations of the  $\{\Delta_{ij}\}$  that are maximally uncorrelated with one another, while the corresponding eigenvalues would indicate the expected variance on a measurement of each linear combination. Sorting these “principal components” (PCs) by eigenvalue, we can then hope to identify a minimal subset of them that will suffice for a particular analysis (in our case, providing sufficient constraints on a range of baryonic models). Such a procedure has previously been applied, for example, to the reconstruction of a time-dependent equation of state for dark energy (Huterer & Starkman 2003), and of the primordial spectrum of curvature perturbations from inflation (Kadota et al. 2005; Dvorkin & Hu 2010).

For a given survey, the covariance matrix of the  $\{\Delta_{ij}\}$  (along with other parameters related to the background cosmology and systematics) can be approximated by the inverse of the appropriate Fisher matrix. Recall that for a given vector of observed data points  $\mathbf{d}$  with associated covariance  $\mathbf{Cov}$ , along with a given model with parameters  $\{p_a\}$ , the Fisher matrix is given by<sup>5</sup>

$$F_{ab} = \frac{\partial \mathbf{d}^T}{\partial p_a} \mathbf{Cov}^{-1} \frac{\partial \mathbf{d}}{\partial p_b} + \frac{\delta^{ab}}{\sigma_a^2} , \quad (2)$$

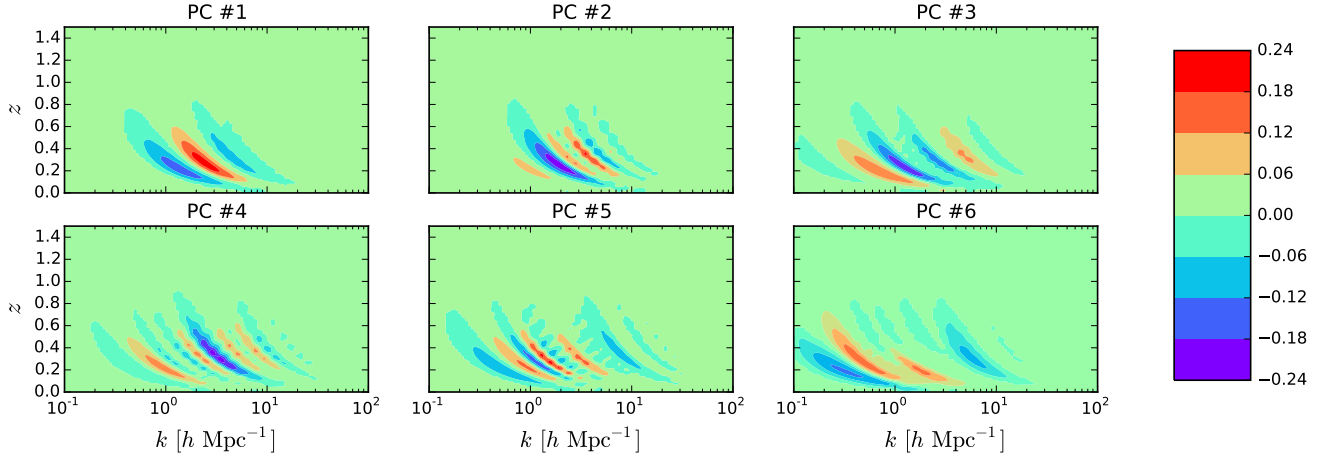
where the last term incorporates optional Gaussian priors on each parameter  $p_a$ . Upon inverting  $F_{ab}$  and then diagonalizing the submatrix corresponding to the  $\{\Delta_{ij}\}$ , we obtain a new set of parameters  $\alpha_a$ , defined by

$$\alpha_a = \sum_{\mu} \beta_{a\mu} \Delta_{\mu} , \quad (3)$$

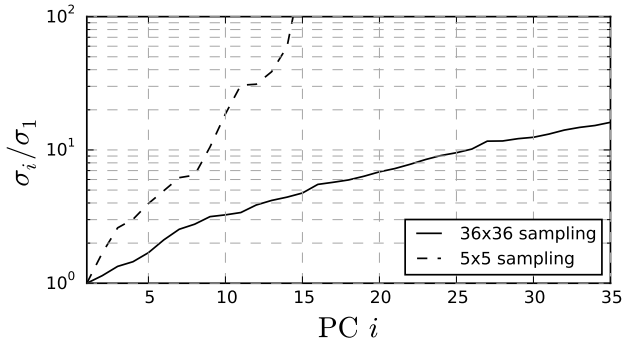
where the  $(i, j)$  indices on  $\Delta$  have been compressed into a single index  $\mu$  for brevity. We use conventions where the  $\{\beta_a\}$  are orthonormal vectors, i.e.  $\sum_{\mu} \beta_{a\mu} \beta_{b\mu} = \delta_{ab}$ . We will denote the eigenvalues (forecast variances of each  $\alpha_a$ ) by  $\sigma_a^2$ .

To visualize the results of this procedure, in Fig. 1 we show the 6 PCs whose amplitudes will be best constrained by (i.e. have the lowest variance when constrained using) measurements of the cosmic shear two-point function from DES Y5, where the assumed properties and treatment of systematics for DES Y5 will be described in Secs. 3.1-3.2. These modes were constructed from a set of  $\{\Delta_{ij}\}$  defined in a  $36 \times 36$  grid in  $\log(k/[h \text{ Mpc}^{-1}])$  and  $z$ , bounded by the edges of each panel in Fig. 1. The lower bound for  $k$ ,  $0.1 h \text{ Mpc}^{-1}$ , has specifically been chosen to exclude linear scales, while the other boundaries have been chosen to encompass the range of wavenumbers and redshifts where DES is expected to be most informative (assuming measurements of  $\xi_{\pm}$  to  $\theta_{\min} = 2'$ ). Similar considerations apply to our forecasts for other surveys.

<sup>5</sup> The Fisher matrix will also contain terms that include the derivative of  $\mathbf{Cov}$  with respect to the model parameters, but we will ignore those terms because they will generally be subdominant for a large-area survey (e.g. Eifler et al. 2009).

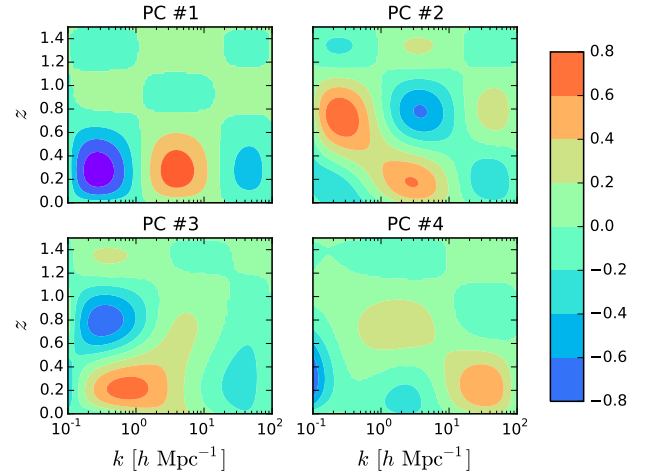


**Figure 1.** The six principal components of  $\Delta(k, z) \equiv P(k, z)/P_{\text{DM-only}}(k, z) - 1$  that will be best constrained by measurements of the cosmic shear correlation functions  $\xi_{\pm}$  from DES Y5, as determined by the procedure in Sec. 2. (See Sec. 3 for our assumptions about the details of these measurements.) The shapes of these principal components correspond to curves in the  $(k, z)$  plane that contribute to multipoles at different  $\ell$  values.



**Figure 2.** Expected one-sigma uncertainties  $\sigma_i$  on the amplitudes  $\alpha_i$  of the most relevant PCs for DES Y5, normalized to the uncertainty on the best-constrained mode. The solid line corresponds to PCs constructed from a fine sampling of the  $(k, z)$  plane (a  $36 \times 36$  grid in  $\log(k/[h \text{ Mpc}^{-1}])$  and  $z$ ), while the dashed line corresponds to a coarser sampling (a  $5 \times 5$  grid). The expected uncertainties increase rapidly with mode number, implying that a finite number of the lowest modes likely contain the bulk of a survey’s information about the matter power spectrum. Furthermore, this increase is most rapid when the  $(k, z)$  plane is sampled more coarsely, since each sample then acts as a wider “bin” in the  $(k, z)$  plane and thus encapsulates more information. Forecasts for other surveys exhibit the same behavior.

The appearance of these modes in Fig. 1 reflects the sensitivity of the data to projections of  $P(k, z)$  that contribute to different angular scales: specifically, a given multipole  $C_\ell$  will probe the matter power spectrum along the curve  $k = \ell/\chi(z)$ , where  $\chi(z)$  is the comoving distance to redshift  $z$  (see Eq. (5) in Sec. 3.1). Measurements of  $\xi_{\pm}$  in a given angular bin will be sensitive to a range of multipoles; it is also possible that sets of angular bins exist that are impacted equally by fractional variations in the power spectrum at a given level. Fig. 1 demonstrates that these angular bins (and their corresponding ranges of multipoles)



**Figure 3.** As Fig. 1, but displaying PCs constructed from a set of  $\{\Delta_{ij}\}$  that sample the  $(k, z)$  plane a factor of  $\sim 5$  more coarsely in both  $k$  and  $z$ . With this coarser sampling, measurements at a greater range of angular scales are important for each PC, with the result that these four PCs contain a much larger fraction of the information to be found in DES than the first four PCs shown in Fig. 1. This type of sampling of the  $(k, z)$  plane is representative of what we will use in the remainder of this work.

are identified and combined together automatically by the PCA procedure.

In Fig. 2, we show the projected constraints on the amplitudes of the most important PCs for DES Y5, normalized to the constraint on the PC with lowest expected variance. The solid line in this figure uses PCs obtained from the same  $36 \times 36$  grid in  $\log(k/[h \text{ Mpc}^{-1}])$  and  $z$  for  $\Delta_{ij}$  that was used to produce Fig. 1, while the dashed line arises from a much coarser  $5 \times 5$  grid in  $\log(k/[h \text{ Mpc}^{-1}])$  and  $z$ . (The four best-constrained PCs corresponding to this coarser sampling are shown in Fig. 3.)

The steep dropoff in constraining power (generically



also seen in forecasts for other surveys, not shown in Fig. 2) indicates that a finite number of PCs will likely contain the majority of the information about the matter power spectrum that we could expect to extract from a given set of shear measurements. The difference between the dashed and solid lines further implies that a coarser sampling of the  $(k, z)$  plane is more efficient in compressing the constraining power into a small set of PCs. Henceforth, in this work we will use a  $5 \times 5$  grid to produce our results, since this is the smallest number of points that will allow us to use cubic splines to interpolate along each axis. Coarser grids paired with different interpolation methods may also be useful, but we will not pursue them in this work.

It is worth mentioning that in the original basis of  $\{\Delta_{ij}\}$  variables, some of these variables may be so poorly constrained that their corresponding elements in the Fisher matrix will cause the matrix to become ill-conditioned, impeding its inversion or diagonalization. This can be ameliorated by incorporating a large Gaussian prior for each variable, which will not affect the results for the most important modes. For our forecasts in this work, a prior of  $\sigma(\Delta_{ij}) = 10^4$  was sufficient to eliminate these numerical issues.

### 3 MOCK DATA AND BARYONIC MODELS

We will demonstrate the performance of the method described in Sec. 2 by carrying out forecasts in which it is applied to a number of mock data sets corresponding to current and upcoming weak lensing surveys, and comparing to what would be obtained from attempting to directly constrain a set of representative models for baryonic effects on the matter power spectrum. In this section, we describe how these mock data sets are constructed, our treatment of systematic effects that could be important for cosmic shear measurements on the relevant scales, and the baryonic models we use for our test set.

#### 3.1 Mock data

For our forecasts, we consider a variety of upcoming surveys (along with the already-completed Dark Energy Survey Science Verification run, DES SV, which will serve as an indicator of the constraining power of currently available data). Table 1 displays the properties we will assume in this work. These properties are based mainly on Krause et al. (2016), with following exceptions: for DES Y5, we have used a slightly more conservative value for the density of sources; for HSC, we have used the sky coverage and a more conservative value of the source density from Osato et al. (2015);

and for Euclid, we have used the density of sources specified in Laureijs et al. (2011). For all surveys, we will assume an intrinsic shape noise of  $\sigma_\epsilon = 0.27$  per component. Where appropriate, throughout the paper we will comment on the sensitivity of our forecasts to these assumptions.

For the data themselves, we will use the real-space two-point shear correlation functions  $\xi_\pm(\theta)$ , given by

$$\xi_\pm^{ij}(\theta) = \frac{1}{2\pi} \int d\ell \ell J_{0/4}(\ell\theta) C_\ell^{ij}, \quad (4)$$

where  $J_{0/4}$  is the Bessel function of order 0 or 4, and  $i, j$  denote the two tomographic redshift bins involved in the correlation. The multipoles of the convergence field,  $C_\ell^{ij}$ , are obtained from the matter power spectrum via

$$C_\ell^{ij} = \frac{9H_0^4\Omega_m^2}{4c^4} \int_0^{\chi_h} d\chi \frac{g^i(\chi)g^j(\chi)}{a^2(\chi)} P\left(\frac{\ell}{f_K(\chi)}, z(\chi)\right), \quad (5)$$

where  $\chi_h$  is the comoving distance to the horizon and  $f_K(\chi)$  is the comoving angular diameter distance (equal to  $\chi$  in a flat universe, which we will assume in this work). The lens efficiency  $g_i(\chi)$  is given by

$$g^i(\chi) = \int_\chi^{\chi_h} d\chi' n_i(\chi') \frac{f_K(\chi' - \chi)}{f_K(\chi')}. \quad (6)$$

Our base forecasts will be performed using 6 measurements of each of  $\xi_+$  and  $\xi_-$  per redshift bin pair, log-spaced between  $\theta_{\min} = 2'$  and  $\theta_{\max} = 300'$ , although we will also investigate the consequences of other choices of  $\theta_{\min}$ .

For the distribution of sources  $n_i(z)$  in each redshift bin, we will use the exact distributions from DES SV (Becker et al. 2015) for the corresponding forecast. For DES Y5 and the other surveys, we employ the following parametrization with parameter values given in Table 1, dividing the total  $n(z)$  into bins with equal numbers of sources:

$$n(z) \propto \Theta(z_{\max} - z) z^\alpha \exp\left[-\left(\frac{z}{z_0}\right)^\beta\right]. \quad (7)$$

We base the source distribution for HSC on Oguri & Takada (2011), while we follow Krause et al. (2016) for the other surveys. We make a conservative choice of 5 redshift bins for DES Y5, while we will assume that the other surveys will be capable of collecting sufficient statistics in 10 redshift bins.

For the covariance of the mock correlation function measurements, we will restrict ourselves to the Gaussian approximation, evaluated as in Joachimi et al. (2008). We refer the reader to that paper for the derivation, and merely quote the final expression we use, which accounts for the difference in our conventions for  $\sigma_\epsilon$ :

$$\begin{aligned} \text{Cov}_G[\xi_\pm^{ij}(\theta_1), \xi_\pm^{kl}(\theta_2)] &= \frac{1}{8\pi^2 f_{\text{sky}}} \int d\ell \ell J_{0/4}(\ell\theta_1) J_{0/4}(\ell\theta_2) \left[ \tilde{C}_{\ell_1}^{ik} \tilde{C}_{\ell_1}^{jl} + \tilde{C}_{\ell_1}^{il} \tilde{C}_{\ell_1}^{jk} - (\delta^{ik} \delta^{jl} + \delta^{il} \delta^{jk}) \frac{\sigma_\epsilon^4}{n^2} \right] \\ &\quad + (\delta^{ik} \delta^{jl} + \delta^{il} \delta^{jk}) \delta_{\theta_1 \theta_2} \delta_{+-} \frac{\sigma_\epsilon^4}{8\pi^2 f_{\text{sky}} \theta_1 \Delta \theta_1 n^2}, \end{aligned} \quad (8)$$

where  $\delta_{+-} = 1$  for  $\text{Cov}_G[\xi_+, \xi_+]$  or  $\text{Cov}_G[\xi_-, \xi_-]$  and 0 otherwise, and  $n$  is the angular density of source galaxies in

**Table 1.** Assumed properties of surveys considered in this work

	Sky coverage (deg <sup>2</sup> )	Source density, $\bar{n}$ (gals/arcmin <sup>2</sup> )	Number of redshift bins	Source distribution, $n(z)$
DES SV	139	5.7	3	From (Becker et al. 2015)
DES Y5	5000	8	5	Eq. (7): $z_{\max} = 2.0$ , $\alpha = 1.75$ , $z_0 = 0.25$ , $\beta = 1.0$
HSC	1400	20	10	Eq. (7): $z_{\max} = 3.0$ , $\alpha = 2.0$ , $z_0 = 0.35$ , $\beta = 1.0$
LSST	18000	26	10	Eq. (7): $z_{\max} = 3.5$ , $\alpha = 1.25$ , $z_0 = 0.50$ , $\beta = 1.0$
Euclid	15000	30	10	Eq. (7): $z_{\max} = 2.5$ , $\alpha = 1.25$ , $z_0 = 0.35$ , $\beta = 1.0$
WFIRST	2200	45	10	Eq. (7): $z_{\max} = 4.0$ , $\alpha = 1.25$ , $z_0 = 0.60$ , $\beta = 1.0$

each redshift bin (in our forecasts,  $n$  is the same for each bin in a given survey). In Eq. (8), we have defined

$$\tilde{C}_\ell^{ij} = C_\ell^{ij} + \delta_\ell^{ij} \frac{\sigma_\epsilon^2}{n}. \quad (9)$$

For a subset of our forecasts, we have also implemented the leading non-Gaussian terms in the covariance (the trispectrum of the convergence, along with a halo sample variance term that accounts for the influence of modes beyond the survey window), using the halo model treatment described in Eifler et al. (2015), to which we refer the reader for details. We find that the inclusion of these terms has negligible impact on our results, justifying our use of the Gaussian approximation for the remaining forecasts.

For our fiducial cosmology, we use the following best-fit parameters from Planck Collaboration et al. (2015):  $\Omega_m = 0.316$ ,  $\Omega_b = 0.0491$ ,  $h_0 = 0.673$ ,  $\sigma_8 = 0.83$ ,  $n_s = 0.965$ ,  $w = -1.0$ , and  $\tau = 0.078$ . In our forecasts, we allow these parameters to vary within the one-sigma priors determined from Planck. Previous work (Natarajan et al. 2014; Harnois-Déraps et al. 2015) has shown that it is important to consider the effect of massive neutrinos jointly with other effects on the small-scale matter power spectrum, and therefore we also marginalize over the sum of neutrino masses in our forecasts, assuming a fiducial value of  $\Sigma m_\nu = 0.06\text{eV}$  (for simplicity, we only assume a single massive neutrino species).

For each survey, we compute a mock data vector and mock covariance matrix with the properties and fiducial cosmology described above. We then use these to compute a Fisher matrix, where the set of varied parameters consists of the cosmological parameters in the previous paragraph, neutrino mass, the grid of 25  $P(k, z)$  values described in Sec. 2, and the systematics parameters listed in the next section. The maximum redshift for the power spectrum grid is set roughly to the maximum redshift for each survey, while the maximum wavenumber is set based on the value  $\theta_{\min}$ .<sup>6</sup> We then invert the Fisher matrix and diagonalize the submatrix corresponding to the  $P(k, z)$  samples; the eigenvectors (PCs) and eigenvalues  $\{\sigma_a^2\}$  are then used to obtain the results in the remainder of the paper.

We use CAMB (Lewis et al. 2000) to calculate the linear matter spectrum, and Halofit (Smith et al. 2003; Takahashi

et al. 2012) as an approximation for the nonlinear corrections from gravitational evolution. Furthermore, we use the modified Halofit prescription from Bird et al. (2012) to account for the effect of massive neutrinos. (In Sec. 3.2, we describe how we account for possible uncertainties in the modeling of the nonlinear power spectrum.) Most numerical computations are performed using CosmoSIS (Zuntz et al. 2015), with the exception of the mock covariances, for which we use CosmoLike (Krause & Eifler 2016).

### 3.2 Treatment of systematic errors

We have attempted to include a selection of possible systematic errors that will likely be relevant for cosmic shear measurements on the scales we are interested in. Our implementation is as follows:

(i) **Shear calibration errors:** Following the treatment of the DES Science Verification Data in The Dark Energy Survey Collaboration et al. (2015), we assign a single free multiplicative parameter  $m_i$  to each redshift bin, so that  $C_\ell^{ij} \rightarrow (1 + m_i)(1 + m_j)C_\ell^{ij}$ , and marginalize over each of these parameters, using a Gaussian prior of width 0.05 centered at zero. This choice is extensively discussed and motivated in Jarvis et al. (2015) in the context of DES SV, but the general features of that discussion are likely to apply to shear calibration issues in other surveys as well.

(ii) **Photometric redshift uncertainties:** For each redshift bin, we allow for a uniform translation  $\delta z_i$  of the entire source distribution within that bin, such that  $n_i(z) \rightarrow n_i(z - \delta z_i)$ . We marginalize over each  $\delta z_i$  with a prior of 0.05, which was found to reflect the uncertainty in the redshift distributions from DES SV (Bonnett et al. 2015); while this value is much larger than the uncertainty goal for photo- $z$ 's from future (e.g. Stage III) observations, to be conservative, we use it for all surveys we consider.

(iii) **Intrinsic alignments:** We assume the non-linear tidal alignment model of Bridle & King (2007), with free overall amplitude  $A_{\text{IA}}$ , assigned a wide prior of  $\sigma(A_{\text{IA}}) = 5$  and marginalized over in our analysis. We discuss this choice further in Sec. 5.3.

(iv) **Uncertainty in power spectrum modeling:** We treat this uncertainty by allowing for two constant fractional shifts in the value of the “DM+neutrinos” power spectrum, above and below some transition wavenumber  $k_{\text{tr}}$ , with exponentials smoothly connecting the two regimes:

$$P(k, z) \rightarrow P(k, z) \left[ 1 + S_P^{(\text{low})} \exp(-k^2/k_{\text{tr}}^2) + S_P^{(\text{high})} (1 - \exp[-k^2/k_{\text{tr}}^2]) \right]. \quad (10)$$

<sup>6</sup> Specifically, we use  $k_{\max} = \{50, 100, 400\} h \text{Mpc}^{-1}$  for  $\theta_{\min} = \{4', 2', 0.5'\}$ . These  $k_{\max}$  values we chosen to obtain convergence of the final forecast results with respect to small changes in  $k_{\max}$ . Therefore, the results are not strongly sensitive to the values of the power spectrum precisely at  $k_{\max}$ , but are only sensitive up to a wavenumber around a factor of a few smaller.

We fix  $k_{\text{tr}} = 0.5 h \text{ Mpc}^{-1}$ , and marginalize over the two amplitudes  $S_P^{(\text{low})}$  and  $S_P^{(\text{high})}$ , with Gaussian priors of 0.02 and 0.05 respectively. These priors are roughly based on conservative estimates of the accuracy of currently-available fitting functions like Halofit or the CosmicEmu emulator (Heitmann et al. 2014). In the future, the combination of improvements in perturbation theory (e.g. Foreman et al. 2016) at large scales and more ambitious emulation programs or approaches to nonlinear modeling at smaller scales will likely decrease these errorbars dramatically. We explore the effect of varying the prior on  $S_P^{(\text{high})}$  in Sec. 5.3 (our results are insensitive to the prior on  $S_P^{(\text{low})}$ ).

(v) **Other small-scale shear systematics:** There exist several other systematics effects that could become relevant at high multipoles, or equivalently small ( $\sim$ arcminute) angular scales. Examples of these include higher-order terms in the expansion of the “reduced shear”  $\gamma/(1 - \kappa)$ , which is what is actually observed (rather than the pure shear  $\gamma$  itself) (Shapiro 2009; Krause & Hirata 2010); *lensing bias*, the effect of lensing on observed properties of galaxies, which can lead to biased selection in regions of large magnification (Schmidt et al. 2009; Krause & Hirata 2010); and environment-dependent selection biases, arising for example from excluding blended sources (which typically occur in regions of high convergence) from a shear catalogue (Hartlap et al. 2011; MacCrann 2016; MacCrann et al. 2016).

Fortunately, a simple power-law model,  $\Delta C_\ell / C_\ell \propto \ell^p$  serves to mimic the impact of most of these effects quite closely, with  $p \approx 0.5$  providing a good match to examples of explicit calculations found in the literature (Shapiro 2009; Schmidt et al. 2009). Therefore, we roll these effects into the following phenomenological model, letting  $S_{\text{shear}}$  be the overall amplitude of these effects at  $\ell = 10^4$ :

$$C_\ell^{ij} \rightarrow C_\ell^{ij} \left[ 1 + S_{\text{shear}} \left( \frac{\ell}{10^4} \right)^{0.5} \right]. \quad (11)$$

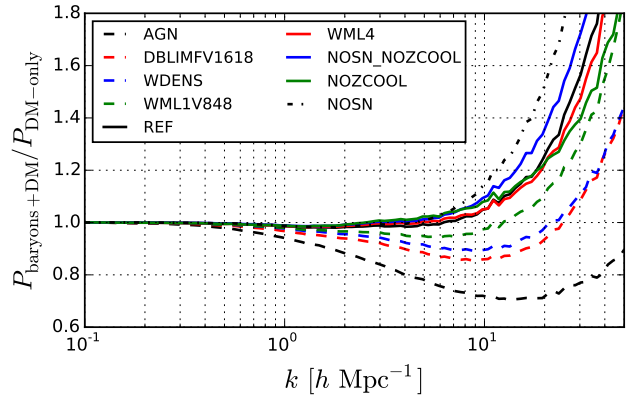
Such a model was also used in Huterer et al. (2006), who pointed out that other additive systematics such as anisotropies in a telescope’s point-spread function will likely have a similar impact on the angular power spectrum.

We marginalize over the amplitude  $S_{\text{shear}}$  with a prior of 0.05, which is roughly the level at which these effects are expected to contaminate near-term shear measurements (MacCrann 2016; MacCrann et al. 2016). We investigate how our results depend on the prior on  $S_{\text{shear}}$  in Sec. 5.3.

In summary, we marginalize over  $(2n_{\text{bins}} + 4)$  systematic parameters, along with the sum of neutrino masses.

### 3.3 Models for baryonic effects

To represent a range for the types of baryonic effects on the matter power spectrum that can be expected within reasonable scenarios, we consider the 9 unique models that compose the OWLS suite of hydrodynamic simulations (Schaye et al. 2010; Van Daalen et al. 2011). A brief summary of the physical effects that are included in each model can be found in Table 1 of Van Daalen et al. (2011). These simulations were carried out using a modified version of Gadget-3, in periodic boxes with side length  $100 h^{-1} \text{ Mpc}$  using 512<sup>3</sup> dark matter particles and an equal number of baryonic particles. Each simulation (including one that contains only dark



**Figure 4.** Ratio of total (baryons+DM) power spectra over DM-only power spectrum at  $z = 0$  for each baryonic model described in Sec. 3.3.

matter) was started with the same initial conditions, so that the ratio  $P_{\text{baryons+DM}}/P_{\text{DM-only}}$  can be obtained without the sample variance inherent in the initial perturbations<sup>7</sup>.

These ratios are shown in Fig. 4 at  $z = 0$  for each of the 9 models. These models can be phenomenologically classified into two groups:

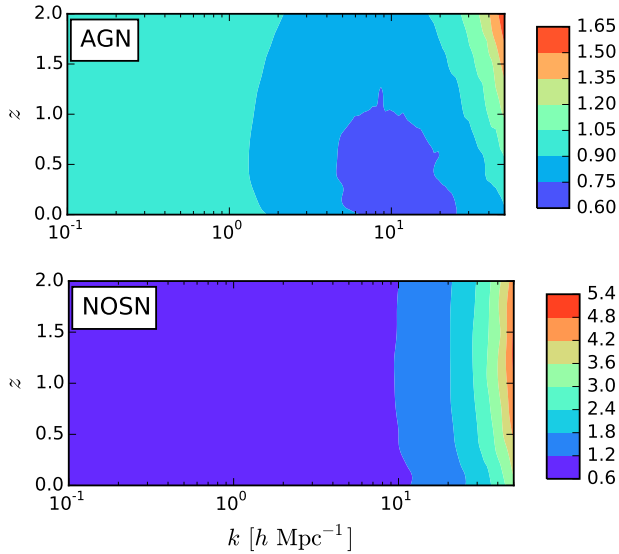
(i) *AGN*, *DBLIMFV1618*, *WDENS*, *WML1V848*: These include strong feedback processes that drive gas out of galaxies, smoothing the total density field and thereby suppressing the power spectrum for  $k \lesssim 10 h \text{ Mpc}^{-1}$ .

(ii) *REF*, *WML4*, *NOSN\_NOZCOOL*, *NOZCOOL*, *NOSN*: These contain relatively weak feedback (or none at all), with the dominant effect on the power spectrum instead being an enhancement at  $k \gtrsim 5 h \text{ Mpc}^{-1}$  arising from gas cooling that increases the central density in halos.

In Fig. 5, we show contour plots of the ratio  $P_{\text{baryons+DM}}/P_{\text{DM-only}}$  for the two models at the extremes of this classification: *AGN*, which includes feedback from active galactic nuclei in addition to supernovae, and *NOSN*, which does not include any feedback mechanisms but includes radiative cooling.

Of the OWLS models, *AGN* can be thought of as the most realistic, since the *AGN* feedback effects that it includes result in much better agreement with observations of low-redshift groups of galaxies than any of the other models (see Schaye et al. 2010; Van Daalen et al. 2011; Viola et al. 2015 for details). Some others, such as those that neglect radiative cooling processes known to be important in matching galaxy-scale observations, are decidedly unre-

<sup>7</sup> Of course, these ratios will likely have some dependence on the background cosmology, but since we are only using these models as representative of the possible range of baryonic effects on the matter power spectrum, we will ignore this cosmology-dependence in this work. Furthermore, the ratio  $P_{\text{baryons+DM}}/P_{\text{DM-only}}$  will itself have nonzero sample variance, by virtue of the fact that it has been measured from a finite spatial volume. We will not attempt to quantify this sample variance in this work, but merely note that future precision comparisons with hydrodynamical simulations will eventually need to take this into account.



**Figure 5.** As Fig. 4, but plotted as a function of wavenumber and redshift, and only showing the two most extreme models: one (*AGN*) in which the combined effects of AGN and supernova feedback lead to a strong suppression of the power spectrum at relatively large scales, and another (*NOSN*) that neglects all feedback processes but contains gas cooling that enhances the power spectrum at small scales.

alistic. Nevertheless we still find it useful test to test our principal-component method on the full range of scenarios.

#### 4 VALIDATION OF METHOD

The strategy of identifying the best-constrained principal components of  $P(k, z)$ , as described in Sec. 2, will only be useful if these principal components can also provide a reasonable amount of information about the influence of baryonic effects on the power spectrum. In this section, we quantify the information content of the best-constrained PCs with respect to our chosen set of baryonic models, and demonstrate that a small number of these PCs provides sufficient information about the full set of models to lead to meaningful constraints.

We first define the following parametrization for each baryonic model separately, based around a free amplitude  $A_{\text{bar}}$  that interpolates between the DM-only case ( $A_{\text{bar}} = 0$ ) and the exact DM+baryons power spectrum ( $A_{\text{bar}} = 1$ ):

$$P(k, z) = P_{\text{DM-only}}(k, z) \times \left[ 1 + A_{\text{bar}} \left( \frac{P_{\text{DM+baryons}}(k, z)}{P_{\text{DM-only}}(k, z)} - 1 \right) \right]. \quad (12)$$

If we were only interested in a single baryonic model, we could imagine using Eq. (12) as our model for baryonic effects, and attempting to constrain  $A_{\text{bar}}$  along with the other cosmological and systematics parameters. Alternatively, we can parametrize baryonic effects using the amplitudes  $\{\alpha_a\}$  of the  $N$  most important PCs. To do so, we can solve Eq. (3) for  $\Delta_\mu$  at each  $\mu \equiv (k_i, z_j)$ , using only up to mode  $N$  on the right hand side, and then smoothly interpolate between these discrete values of  $\Delta_\mu$  to construct a function  $\Delta_N(k, z)$ ,

which will depend on the values  $\{\alpha_1, \dots, \alpha_N\}$ . The power spectrum is then given by

$$P(k, z) = P_{\text{DM-only}}(k, z) [1 + \Delta_N(k, z)]. \quad (13)$$

An intuitive way to determine how much information the PCs provide about different baryonic models is to relate constraints on the amplitudes  $\{\alpha_a\}$  to a constraint on  $A_{\text{bar}}$ . If we treat each amplitude as providing an independent “measurement” of  $A_{\text{bar}}$  and apply inverse-variance weighting to these “measurements,” we obtain<sup>8</sup>

$$\frac{1}{\sigma^2(A_{\text{bar}})} = \sum_{a=1}^N \frac{1}{\sigma_a^2} \left( \frac{\partial \alpha_a}{\partial A_{\text{bar}}} \right)^2 = \sum_{a=1}^N \frac{(\alpha_a^{(\text{bar})})^2}{\sigma_a^2}, \quad (14)$$

where  $\alpha_a^{(\text{bar})}$  is the projection of the specific baryonic model onto the chosen PCs:

$$\alpha_a^{(\text{bar})} = \sum_{\mu} \beta_{a\mu} \Delta_\mu^{(\text{bar})}. \quad (15)$$

(The second equality in Eq. (14) follows from the fact that  $\Delta_\mu^{(\text{bar})}$ , and therefore  $\alpha_a^{(\text{bar})}$ , is linear in  $A_{\text{bar}}$ .)

When *all* available PCs are used in Eq. (14) to determine  $\sigma(A_{\text{bar}})$ , the result is very close to the output of forecasts in which we constrain  $A_{\text{bar}}$  directly for each baryonic model in our testing set. However, it is sufficient to use a much smaller number of PCs to retain a comparable amount of constraining power. In Fig. 6, we check how this constraining power scales with the number  $N$  of the most important PCs we retain, by plotting the ratio of  $\sigma(A_{\text{bar}})$  obtained from all PCs and that from the first  $N$  PCs. If we wish to retain 90% of the constraining power of the full set of PCs, we may only keep the first 9 PCs in our analysis. This number is an upper bound across all surveys we have investigated, including those not shown in Fig. 6, and can be considered as a validation of the usefulness of the PCA approach.

In the next section, we will perform forecasts for the constraining power of the first 9 PCs of various surveys. We use the same number of PCs for each survey in order to make meaningful comparisons, but in principle, a given survey might require even fewer PCs to achieve meaningful constraints. For example, the top panel of Fig. 6 shows that for cosmic shear measurements comparable to those of DES SV, only the first 6 modes are required to capture 90% of the full amount of information about the impact of baryons of the matter power spectrum. Optimization of the number of PCs in other cases is best left to future, survey-specific analyses.

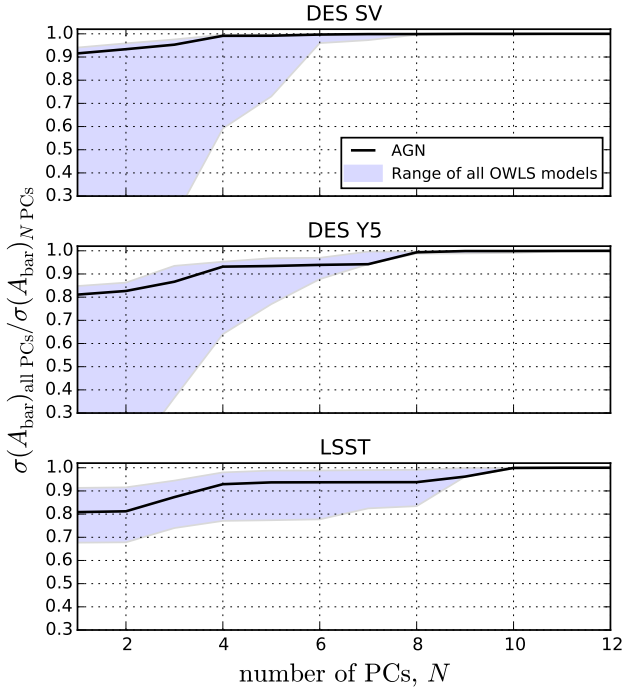
## 5 FORECASTS

### 5.1 Performance of various surveys

In order to quantify the constraining power of different surveys with respect to generic baryonic effects on the matter

<sup>8</sup> The right-hand side of Eq. (14) is also the square of the cumulative signal-to-noise of the first  $N$  PCs, and therefore can alternatively be interpreted without the need of the  $A_{\text{bar}}$  parametrization from Eq. (12).





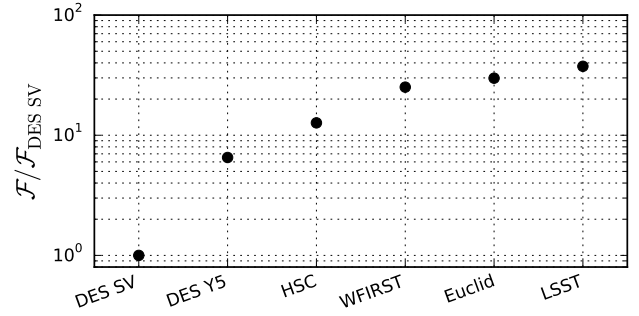
**Figure 6.** Ratio of the effective one-sigma constraint on  $A_{\text{bar}}$  (defined in Eq. (12)),  $\sigma(A_{\text{bar}})$  from using all available PCs with that from using only the  $N$  best-constrained PCs for the three surveys specified above each panel. For each set of modes, the effective constraint on  $A_{\text{bar}}$  is calculated using Eq. (14). The black curve corresponds to the OWLS AGN model, while the blue band represents the range spanned by all OWLS models. This represents the fraction of the full survey’s information that is retained by the first  $N$  PCs of  $P(k, z)$ . Other Stage IV surveys give results comparable to LSST, and therefore we conclude that for all surveys we consider, the first  $\sim 9$  PCs contain 90% of the total constraining power on our test set of baryonic models.

power spectrum, we define a figure of merit  $\mathcal{F}$  as the reciprocal of the geometric mean of the 68% confidence intervals for the most important 9 PCs for each survey:

$$\mathcal{F} \equiv \left[ \prod_{a=1}^9 \sigma_a \right]^{-1/9}. \quad (16)$$

Since the PCs are uncorrelated by construction, the product  $\prod_{a=1}^9 \sigma_a$  is proportional to the volume of the hyperellipsoid corresponding to the joint 68% confidence region of all PCs, but this product differs by many orders of magnitude for different surveys. On the other hand, the quantity in Eq. (16) efficiently captures a kind of average of the constraints on the first 9 modes; from Eq. (14), we then find that the effective constraints on a generic model for baryonic effects (as parametrized by  $\sigma(A_{\text{bar}})$ ) will scale proportionally to  $\mathcal{F}$ .

Fig. 7 presents the figure of merit for each survey we consider in this work, for forecasts using either  $\xi_{\pm}$  measurements down to  $\theta_{\text{min}} = 2'$ , normalized to the figure of merit for DES SV to emphasize the improvement of future shear measurements over currently available data. Fig. 7 indicates that HSC and the Y5 shear measurements from DES will improve on current data by roughly an order of magnitude, while upcoming Stage IV surveys will only improve on that



**Figure 7.** Figure of merit  $\mathcal{F}$  (defined in Eq. (16)) for constraints on baryonic models, normalized to the value for DES SV, for forecasts performed with mock  $\xi_{\pm}$  measurements down to  $\theta_{\text{min}} = 2'$ . We find that both HSC and the Y5 shear measurements of cosmic shear from DES will improve  $\mathcal{F}$  by roughly an order of magnitude, implying a similar potential improvement in our information about baryonic effects on the matter power spectrum, while Stage IV surveys like LSST will improve on that by a further factor of a few.

by a factor of a few. This can be understood intuitively by making use of the following rough scalings:

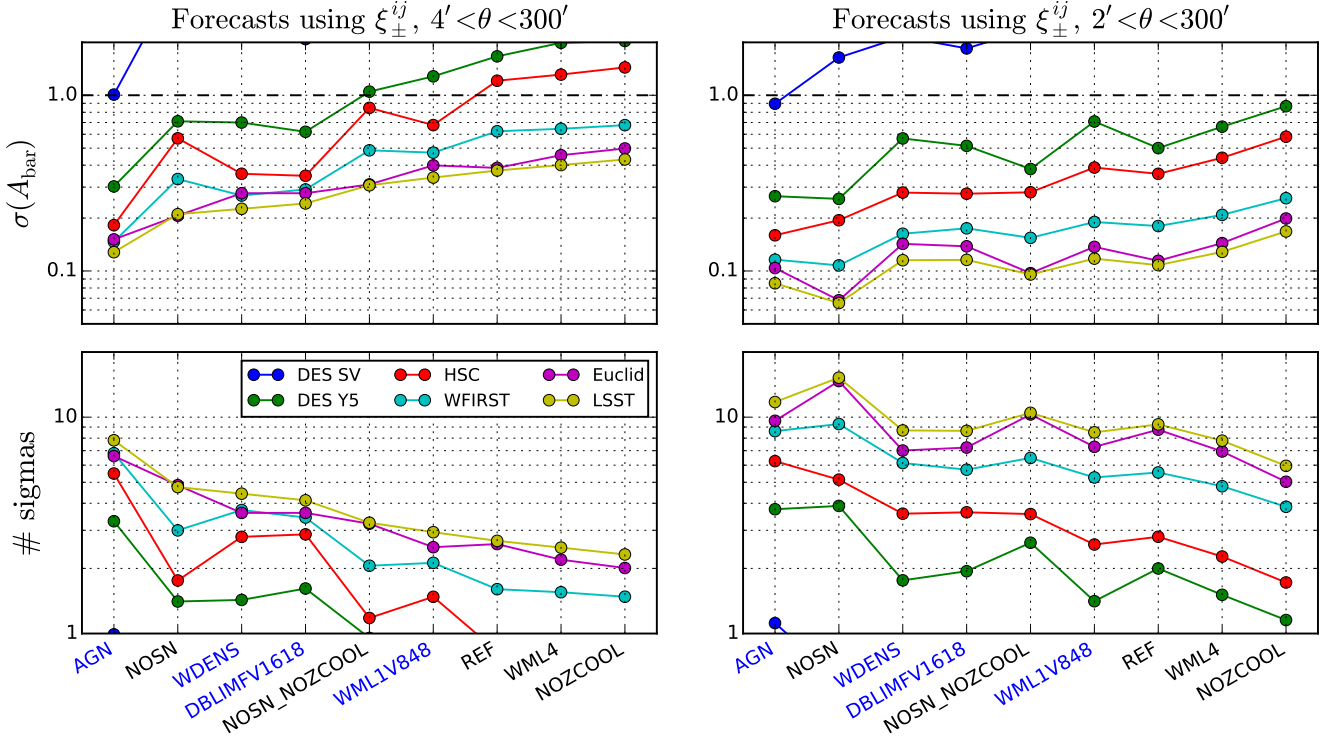
- $\mathcal{F} \propto f_{\text{sky}}^{0.5}$ , since the leading terms in the data covariance matrix scale like  $f_{\text{sky}}^{-1}$ .
- Because of shot noise in the shear correlation function,  $\mathcal{F} \propto (n\sigma_{\epsilon}^{-2})^{0.6}$  (this power has been empirically determined by varying  $n$  in forecasts for various surveys). If shot noise completely dominated the covariance matrix at all scales, we would expect  $\mathcal{F} \propto (n\sigma_{\epsilon}^{-2})^1$ , but the balance between the cosmic variance and shot noise terms in Eq. (9) softens the dependence to what we quote here.
- On average, the figure of merit scales with the number of redshift bins like  $\mathcal{F} \propto n_{\text{bins}}^{0.4}$ . This power is steeper for surveys with broader redshift distributions, since in those cases, narrower redshift bins contain more information than for surveys that primarily probe lower redshifts.

Therefore, it is essentially the improvement in the expected number of galaxies with measured shapes (i.e.  $f_{\text{sky}} \times n$ ) that drives the large improvement of Stage III surveys over current data, as compared to the somewhat more modest improvement afforded by Stage IV surveys.

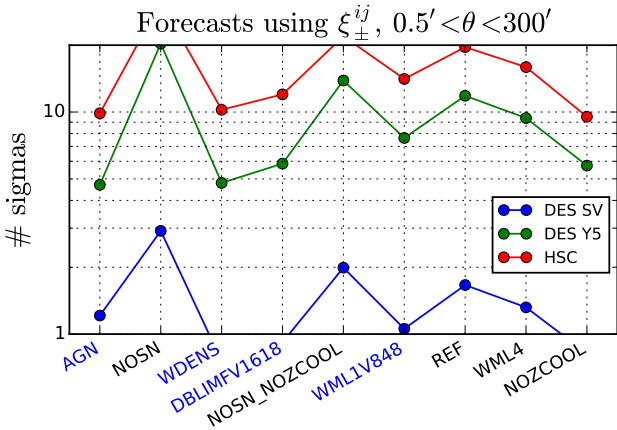
## 5.2 Expected constraints on specific baryonic models

Beyond the model-independent approach of Sec. 5.1, we can also ask how our method can provide information about specific baryonic models when applied to different surveys. Specifically, we can use Eq. (14) to translate the expected constraints on the 9 most important PCs into constraints on  $A_{\text{bar}}$  for each of the models in Sec. 3.3. We can further translate these constraints into the statistical significance at which each model could be ruled out by a given survey (or, alternatively, the significance at which the DM-only power spectrum can be ruled out).

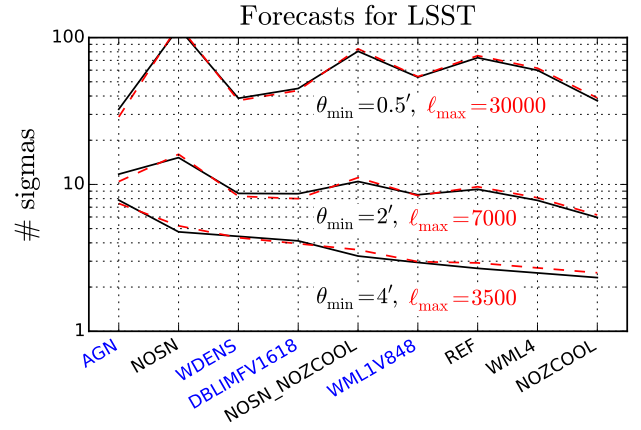
The top left panel of Fig. 8 shows the resulting values for  $\sigma(A_{\text{bar}})$  for each survey and baryonic model, in the case where the shear correlation functions  $\xi_{\pm}^{ij}$  have been measured



**Figure 8.** *Top panels:* Expected one-sigma constraints on  $A_{\text{bar}}$  (see Eq. (12)) for each survey and baryonic model we consider, when the first 9 PCs for each survey are used in the analysis. *Bottom panels:* Statistical significance at which each survey could rule out each model, again using the PCA approach with the amplitudes of the first 9 PCs included as free parameters in the analysis. Using measurements of  $\xi_{\pm}^{ij}$  over the range  $4' < \theta < 300'$  (left panels), the constraints are strongest on models including prescriptions for feedback or other effects that affect the power spectrum at relatively large scales, e.g.  $k \lesssim 1 h \text{Mpc}^{-1}$  at  $z = 0$  (blue labels). When  $\theta_{\min} = 2'$  is used instead (right panels), models with large deviations from the DM-only power spectrum on small scales (typically due to cooling effects on the centers of halos; black labels) are much better constrained due to the inclusion of the additional small-scale information.



**Figure 9.** Same as the bottom panels of Fig. 8, but assuming that measurements of  $\xi_{\pm}$  can be used down to  $\theta_{\min} = 0.5'$ . The gain in constraining power is significant (and also robust to the variations in systematics discussed in Sec. 5.3), with all other surveys we have considered (not shown) providing better than  $20\sigma$  constraints across the entire test set of baryonic models. Therefore, the coming generation of cosmic shear measurements could potentially be very informative with regards to the implementation of certain baryonic phenomena in simulations.



**Figure 10.** Same as the bottom panels of Fig. 8, but comparing LSST forecasts using the real-space correlation function  $\xi_{\pm}$  (black solid lines) or the angular power spectrum  $C_{\ell}$  (red dashed lines) with different values of  $\theta_{\min}$  and  $\ell_{\max}$  as indicated in the figure. For each value of  $\theta_{\min}$ , we find that  $\ell_{\max}$  can be chosen to provide very similar constraints on our test set of baryonic models, following the relationship  $\theta_{\min} \simeq 1.3\pi/\ell_{\max}$ .

down to  $\theta_{\min} = 4'$ , while the bottom left panel shows the corresponding statistical significance of these constraints, given in terms of the “number of sigmas” at which the model could be distinguished from the DM-only case. Unsurprisingly, the models that can be best constrained by the most powerful surveys are *AGN*, whose implementation of feedback has dramatic effects at large scales<sup>9</sup>, and *NOSN*, which lacks any feedback processes but implies a large enhancement at small scales due to cool gas altering the cores of halos. Other models with significant feedback or strong winds, such as *WDENS*, are the next-best constrained, in some cases even better than *NOSN* due to shape noise and other factors that reduce the capabilities of some surveys to probe the smallest scales.

When we consider measurements of  $\xi_{\pm}$  down to  $\theta_{\min} = 2'$ , shown in the right panels of Fig. 8, the corresponding increase in small-scale information serves to improve the constraining power by roughly a factor of two in many cases<sup>10</sup>. The improvement is more modest for models where feedback reduces the impact of cooling on the clustering at the smallest scales, and in some cases, the slight shift in the angular coordinates of the measured  $\xi_{\pm}$  data points even causes a very slight degradation in the constraints.

Given the importance of small scales for this analysis, we have also investigated the case where measurements of  $\xi_{\pm}$  can be used down to  $\theta_{\min} = 0.5'$ . The results for DES SV-like or Stage III datasets are shown in Fig. 9; for all other surveys we consider, the constraints on all baryonic models are better than  $20\sigma$ . This conclusion is insensitive to the priors on systematic effects that we will return to in Sec. 5.3, but of course relies on sufficient control of all observational issues that could affect the usability of measurements at these scales. Nevertheless, it seems likely that data from Stage III surveys, or even partial data releases from these surveys, will be informative with respect to certain behaviors of baryons on the scales probed by cosmic shear measurements.

We have presented our results in the framework of distinguishing each model from the DM-only case, but with the high level of statistical significance that will be afforded by future surveys, it is likely that these measurements will be able to distinguish between different scenarios themselves. This can be accomplished by examining the joint posteriors on the amplitudes of each PC, and then comparing the projections of individual models onto these PCs. This procedure can in principle be applied to any model or simulation for which the total matter power spectrum is known, and may also be useful in constraining models containing their own continuous parameters.

Similar conclusions apply to cases where two-point statistics other than the correlation functions  $\xi_{\pm}$  are used in the analysis. In Fig. 10, we compare forecasts for LSST

performed with either  $\xi_{\pm}$  or the angular power spectrum  $C_{\ell}$ , with  $\ell_{\max}$  chosen to reproduce as closely as possible the results from  $\xi_{\pm}$  for each value of  $\theta_{\min}$  we have previously considered. It is to be expected that  $\theta_{\min}$  scales roughly as  $\pi/\ell_{\max}$ , and our explicit comparisons bear out this expectation, with  $\theta_{\min} \simeq 1.3\pi/\ell_{\max}$ . Specifically, we find that the  $\xi_{\pm}$  forecasts with  $\theta_{\min} = \{4', 2', 0.5'\}$  match quite closely with  $C_{\ell}$  forecasts with  $\ell_{\max} = \{3500, 7000, 30000\}$ , respectively.

### 5.3 Impact of systematics at small scales

Of the possible systematics we describe in Sec. 3.2, two of the more uncertain are the accuracy of the modeling of the DM-only power spectrum (including the effects of massive neutrinos) and the level of shear-specific systematics (e.g. lensing bias) that could be important when interpreting measurements at small scales. Our fiducial choice for the main results of the paper has been a 5% uncertainty in the DM-only power spectrum for  $k > 0.5 h \text{ Mpc}^{-1}$  (our  $S_P^{(\text{high})}$  parameter) and a 5% error budget for shear-specific systematics at  $\ell = 10^4$  (our  $S_{\text{shear}}$  parameter), with a profile scaling like  $\Delta C_{\ell} \propto \ell^{0.5}$ .

In Figs. 11 and 12, we display representative variations in our forecasts when the priors on  $S_P^{(\text{high})}$  and  $S_{\text{shear}}$  are varied away from their fiducial values, for the forecasts with  $\theta_{\min} = 4'$  and  $2'$  respectively. For the *AGN* model and others where feedback affects the power spectrum on large scales, the accuracy in the modeling of the DM-only power spectrum is an important determinant of the overall constraining power, provided that the other shear systematics are controlled at a reasonable level. On the other hand, for models for which the constraining power is concentrated at smaller scales, the effects of baryons on the power spectrum at these scales are generally so strong that knowledge of the DM-only power spectrum to within 10% will be sufficient to detect their presence from shear measurements alone.

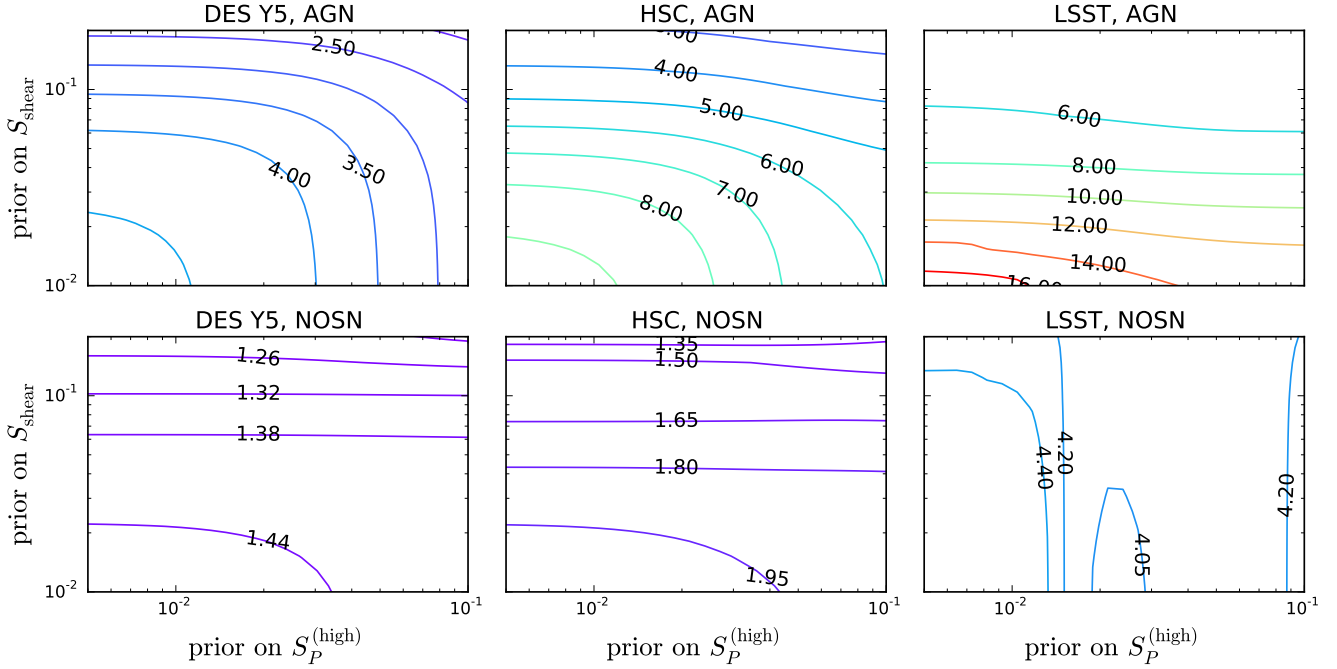
Essentially the same conclusion holds for the prior on  $S_{\text{shear}}$ : it has a relatively smaller impact on cooling-dominated models than on feedback-dominated scenarios. The increased precision of Stage IV surveys like LSST actually translates into *less* stringent requirements on the modeling of the DM-only power spectrum for *AGN*-like models, although the main science goals of these surveys (e.g. searches for new cosmological physics) will of course impose their own stricter requirements. In fact, for cooling-dominated models, we find that the constraints from Stage IV surveys are fairly insensitive to the systematics we vary in this section; measurements with  $\theta_{\min} \lesssim 3'$  will be capable of distinguishing our entire test set of baryonic models from the DM-only case, even with  $\sim 10\%$  uncertainties in the theoretical modeling on the relevant scales.

## 6 CONCLUSIONS

In this paper, we have assessed the ability of the two-point statistics of cosmic shear to act as a probe of the physics of galaxy formation, which is expected to impact the matter power spectrum. We have done so using a method that parameterizes the effects of baryons on the matter power spectrum in terms of a set of principal components (PCs). For a given survey, the method automatically classifies these

<sup>9</sup> Other more recent hydrodynamical simulations, such as Illustris (Vogelsberger et al. 2014), have even stronger effects on the power spectrum at large scales, and therefore can likely be constrained at least at the level of the *AGN* model we examine here.

<sup>10</sup> We have also checked the case where  $\theta_{\min}$  is set to  $2'$  for  $\xi_{+}$  and  $\theta_{\min} = 10'$  for  $\xi_{-}$ , since different scale cuts are sometimes used for  $\xi_{+}$  and  $\xi_{-}$  due to the increased vulnerability of  $\xi_{-}$  to small-systematics as compared to the  $\xi_{+}$ . The results in this case are very similar to what we obtain using  $\theta_{\min} = 4'$  for both  $\xi_{+}$  and  $\xi_{-}$ .



**Figure 11.** For  $\xi_{\pm}$  measurements with  $\theta_{\min} = 4'$ , each panel shows the statistical significance (in “number of sigmas”, as in Figs. 8–10) at which the specified baryonic model can be ruled out by the specified survey, as a function of the assumed level of two types of systematic uncertainty in the theoretical modeling. The parameter  $S_{(\text{shear})}$  denotes the uncertainty in the handling of shear-specific systematics (such as source-lens clustering) in the modeling of the shear correlation function, assumed to scale like  $\Delta C_{\ell} \propto \ell^{0.5}$ , while  $S_P^{(\text{high})}$  denotes the uncertainty in the DM-only power spectrum for  $k < 0.5 h \text{ Mpc}^{-1}$  (see Sec. 3.2 for details). Accuracy in both aspects of the modeling is more important for AGN-like models in which feedback affects the power spectrum on relatively large scales, while for cooling-dominated models. Meanwhile, Stage IV surveys will be sufficiently powerful that relatively large ( $\sim 10\%$ ) uncertainties in the modeling will not appreciably degrade the constraints.

PCs in terms of the expected constraints on each one, allowing for poorly-constrained PCs to be discarded from an analysis without a significant loss of information. These PCs are determined by the properties of a given survey, and are independent of the implementations of baryonic physics one might be interested in testing.

We have performed forecasts for the application of this method to a variety of surveys, ranging from those that have already been completed (the Dark Energy Survey Science Verification run) to Stage III (DES Y5, HSC) and Stage IV (WFIRST, LSST, Euclid) surveys. We have used the OWLS simulation suite, composed of nine different implementations of baryonic effects plus a reference dark matter-only run, to test how the (model-independent) output of the method translates into constraints on a range of specific models.

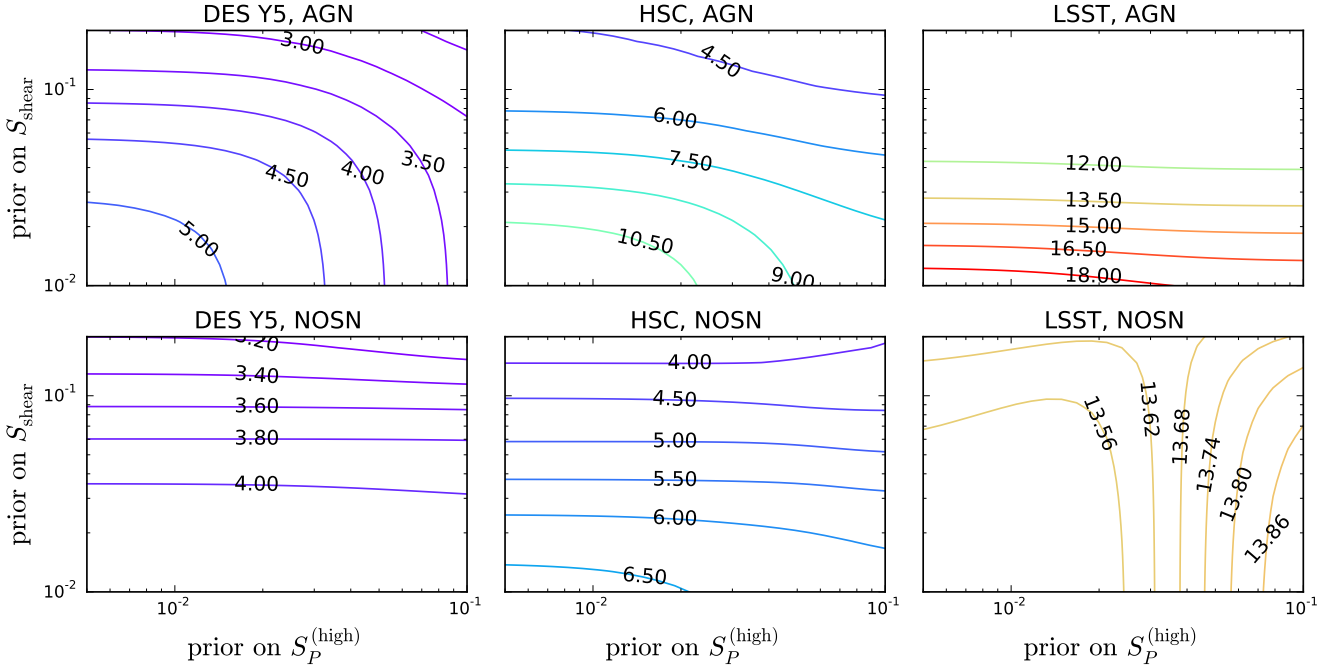
We find that surveys operating now and planned in the near future have significant power to constrain the impact of galaxy formation on the matter power spectrum, with the potential to strongly constrain or rule out a majority of the models considered here. (Indeed, in Harnois-Déraps et al. 2015, mild constraints were already obtained from CFHTLenS data, although using a different method than what we present here.) We emphasize that cosmic shear is a very clean signal to model: the connection between the observed correlations of galaxy shapes and the underlying matter clustering is completely determined by a weak-field calculation in general relativity, and does entail any of the

astrophysical calibration or selection issues involved in many other probes of galaxy formation physics.

Our main results may be summarized as follows:

- For all surveys we consider, we find that 90% of the constraining power with respect to baryonic effects on the matter power spectrum is encapsulated in no more than nine PCs, implying a likelihood analysis that makes use of this PCA method would need to add no more than nine additional parameters.
- From these nine PCs, one can define a figure of merit as the reciprocal of the geometric mean of the expected one-sigma constraints on the corresponding amplitudes. By comparing this figure of merit for different surveys, we find that the constraints from Stage III surveys will improve on those from currently available shear measurements by roughly an order of magnitude, while Stage IV surveys will provide further improvements of a factor of a few. We argue that it is chiefly the number of galaxies with measured shapes that drives this improvement.
- We find that the ultimate power of our method to constrain baryonic effects depends strongly on the minimum scale  $\theta_{\min}$  at which measurements of the shear correlation functions  $\xi_{\pm}$  can be used. In particular, if  $\theta_{\min}$  can be pushed down to less than one arcminute, Stage III surveys should be able to rule out (at more than five-sigma confidence) the majority of the baryonic models we consider. We reach this conclusion after marginalizing over uncertainties in neutrino





**Figure 12.** As Fig. 11, but corresponding to  $\xi_{\pm}$  measurements with  $\theta_{\min} = 2'$  instead of  $4'$ .

mass, photometric redshift distributions, shear calibration, and theoretical modeling of the power spectrum.

In our analysis, we have assumed that the background cosmology has already been mostly fixed by other measurements, such as temperature fluctuations in the cosmic microwave background. For our fiducial cosmology, we have chosen the latest best-fit model from Planck (Planck Collaboration et al. 2015), marginalizing over their one-sigma uncertainties. In applying our method to data, care must be taken in choosing a fiducial cosmology that is consistent with the dataset being employed (accounting, for example, for the current tensions between cosmological constraints from weak lensing and the CMB), to avoid a wrong choice of cosmology biasing the constraints on baryonic effects. While we have not investigated the size of this possible bias in detail, it could easily be investigated in a future analysis simply by varying the choice of cosmology and examining the resulting change in the constraints on the PCs.

The primary product of a likelihood analysis making use of this method would be posterior constraints on the amplitudes of the most important PCs for that survey, with the PCs determined by a similar Fisher matrix procedure to what we have used here. The power spectrum from a given baryonic model can then be projected onto these PCs, and the resulting amplitudes can be compared with the constraints, without the need for a separate likelihood analysis corresponding to every model of interest. In this way, the results of a likelihood analysis including the PCs will act as a resource that can be used to assess the realism of currently-existing models or hydrodynamical simulations, and could also be used to guide the development of new models or simulations.

Furthermore, there currently exist models for baryonic

effects on clustering that contain parameters connected to specific physical processes, such as halo expansion or gas ejection (e.g. Mead et al. 2015; Schneider & Teyssier 2015). Constraints on the PCs from our method could also be mapped onto constraints on such parameters, and could therefore help to inform our knowledge about the related processes, in a more detailed manner than a wholesale acceptance or rejection of an individual model.

Finally, we note that even stronger constraints can likely be obtained by performing this kind of analysis on the combination of cosmic shear and other cosmological statistics. These include, for example, the cluster-mass correlation function, cluster Sunyaev-Zeldovich profiles, and galaxy-galaxy lensing. These may have significant additional power but introduce extra complexity in the modeling due for example to the need for a model for the cluster-halo connection in the first two cases and a bias model in the third case. Regardless, we have shown that, if sub-arcminute measurements can be robustly made, the connection between cosmic shear and the overall clustering of matter can be exploited to turn cosmic shear into an important probe of the physics of galaxy formation.

## ACKNOWLEDGEMENTS

This work received partial support from the U.S. Department of Energy under contract number DE-AC02-76SF00515. We thank Tim Eifler, Andrew Hearin, Dragan Huterer, Elisabeth Krause, Niall MacCrann, and Michael Troxel for useful discussions and comments. We also thank Elisabeth Krause for providing code for computing the covariance matrices used in this work, and Stuart Marshall and the SLAC computing staff for essential computing support.

This work has made extensive use of results from the OWLS simulation suite, and we thank the authors for making these data available from <http://vd11.strw.leidenuniv.nl/>. This work also made use of the CosmoSIS software package, available from <http://bitbucket.org/joezuntz/cosmosis>. We thank Joe Zuntz and the CosmoSIS team for making this code available.

## REFERENCES

- Albrecht A., et al., 2009, [arXiv:0901.0721](#)
- Bacon D. J., Refregier A. R., Ellis R. S., 2000, *MNRAS*, **318**, 625, [arXiv:astro-ph/0003008](#)
- Becker M. R., et al., 2015, [arXiv:1507.05598](#)
- Bernstein G. M., 2009, *ApJ*, **695**, 652, [arXiv:0808.3400](#)
- Bielefeld J., Huterer D., Linder E. V., 2015, *J. Cosmology Astropart. Phys.*, **5**, 023, [arXiv:1411.3725](#)
- Bird S., Viel M., Haehnelt M. G., 2012, *MNRAS*, **420**, 2551, [arXiv:1109.4416](#)
- Bonnett C., et al., 2015, [arXiv:1507.05909](#)
- Bridle S., King L., 2007, *New Journal of Physics*, **9**, 444, [arXiv:0705.0166](#)
- Dvorkin C., Hu W., 2010, *Phys. Rev. D*, **82**, 043513, [arXiv:1007.0215](#)
- Eifler T., Schneider P., Hartlap J., 2009, *A&A*, **502**, 721, [arXiv:0810.4254](#)
- Eifler T., Krause E., Dodelson S., Zentner A. R., Hearin A. P., Gnedin N. Y., 2015, *MNRAS*, **454**, 2451, [arXiv:1405.7423](#)
- Foreman S., Perrier H., Senatore L., 2016, *J. Cosmology Astropart. Phys.*, **5**, 027, [arXiv:1507.05326](#)
- Harnois-Déraps J., van Waerbeke L., Viola M., Heymans C., 2015, *MNRAS*, **450**, 1212, [arXiv:1407.4301](#)
- Hartlap J., Hilbert S., Schneider P., Hildebrandt H., 2011, *A&A*, **528**, A51, [arXiv:1010.0010](#)
- Hearin A. P., Zentner A. R., Ma Z., 2012, *J. Cosmology Astropart. Phys.*, **4**, 034, [arXiv:1111.0052](#)
- Heitmann K., Lawrence E., Kwan J., Habib S., Higdon D., 2014, *ApJ*, **780**, 111, [arXiv:1304.7849](#)
- Heymans C., et al., 2013, *MNRAS*, **432**, 2433, [arXiv:1303.1808](#)
- Huff E. M., Eifler T., Hirata C. M., Mandelbaum R., Schlegel D., Seljak U., 2014, *MNRAS*, **440**, 1322
- Huterer D., Starkman G., 2003, *Physical Review Letters*, **90**, 031301, [arXiv:astro-ph/0207517](#)
- Huterer D., Takada M., 2005, *Astroparticle Physics*, **23**, 369, [arXiv:astro-ph/0412142](#)
- Huterer D., White M., 2005, *Phys. Rev. D*, **72**, 043002, [arXiv:astro-ph/0501451](#)
- Huterer D., Takada M., Bernstein G., Jain B., 2006, *MNRAS*, **366**, 101, [arXiv:astro-ph/0506030](#)
- Jarvis M., et al., 2015, [arXiv:1507.05603](#)
- Jing Y. P., Zhang P., Lin W. P., Gao L., Springel V., 2006, *ApJ*, **640**, L119, [arXiv:astro-ph/0512426](#)
- Joachimi B., Schneider P., Eifler T., 2008, *A&A*, **477**, 43, [arXiv:0708.0387](#)
- Kadota K., Dodelson S., Hu W., Stewart E. D., 2005, *Phys. Rev. D*, **72**, 023510, [arXiv:astro-ph/0505158](#)
- Kaiser N., Wilson G., Luppino G. A., 2000, [arXiv:astro-ph/0003338](#)
- Kitching T. D., Taylor A. N., 2011, *MNRAS*, **416**, 1717, [arXiv:1012.3479](#)
- Kitching T. D., Verde L., Heavens A. F., Jimenez R., 2016, *MNRAS*, **459**, 971, [arXiv:1602.02960](#)
- Krause E., Eifler T., 2016, [arXiv:1601.05779](#)
- Krause E., Hirata C. M., 2010, *A&A*, **523**, A28, [arXiv:0910.3786](#)
- Krause E., Eifler T., Blazek J., 2016, *MNRAS*, **456**, 207, [arXiv:1506.08730](#)
- Kuijken K., et al., 2015, *MNRAS*, **454**, 3500, [arXiv:1507.00738](#)
- Laureijs R., et al., 2011, [arXiv:1110.3193](#)
- Lewandowski M., Perko A., Senatore L., 2015, *J. Cosmology Astropart. Phys.*, **5**, 019, [arXiv:1412.5049](#)
- Lewis A., Challinor A., Lasenby A., 2000, *ApJ*, **538**, 473, [arXiv:astro-ph/9911177](#)
- Lin H., et al., 2012, *ApJ*, **761**, 15, [arXiv:1111.6622](#)
- MacCrann N., 2016, PhD thesis, University of Manchester
- MacCrann N., Zuntz J., Bridle S., Jain B., Becker M. R., 2015, *MNRAS*, **451**, 2877, [arXiv:1408.4742](#)
- MacCrann N., et al., 2016, in preparation
- Mead A. J., Peacock J. A., Heymans C., Joudaki S., Heavens A. F., 2015, *MNRAS*, **454**, 1958, [arXiv:1505.07833](#)
- Mohammed I., Seljak U., 2014, *MNRAS*, **445**, 3382, [arXiv:1407.0060](#)
- Mohammed I., Martizzi D., Teyssier R., Amara A., 2014, *MNRAS*, **440**, 6826
- Natarajan A., Zentner A. R., Battaglia N., Trac H., 2014, *Phys. Rev. D*, **90**, 063516, [arXiv:1405.6205](#)
- Oguri M., Takada M., 2011, *Phys. Rev. D*, **83**, 023008, [arXiv:1010.0744](#)
- Osato K., Shirasaki M., Yoshida N., 2015, *ApJ*, **806**, 186, [arXiv:1501.02055](#)
- Planck Collaboration et al., 2015, [arXiv:1502.01589](#)
- Rudd D. H., Zentner A. R., Kravtsov A. V., 2008, *ApJ*, **672**, 19, [arXiv:astro-ph/0703741](#)
- Samsing J., Linder E. V., Smith T. L., 2012, *Phys. Rev. D*, **86**, 123504, [arXiv:1208.4845](#)
- Schaye J., et al., 2010, *MNRAS*, **402**, 1536, [arXiv:0909.5196](#)
- Schmidt F., Rozo E., Dodelson S., Hui L., Sheldon E., 2009, *ApJ*, **702**, 593, [arXiv:0904.4703](#)
- Schneider A., Teyssier R., 2015, *J. Cosmology Astropart. Phys.*, **12**, 049, [arXiv:1510.06034](#)
- Semboloni E., Hoekstra H., Schaye J., van Daalen M. P., McCarthy I. G., 2011, *MNRAS*, **417**, 2020, [arXiv:1105.1075](#)
- Semboloni E., Hoekstra H., Schaye J., 2013, *MNRAS*, **434**, 148, [arXiv:1210.7303](#)
- Shapiro C., 2009, *ApJ*, **696**, 775, [arXiv:0812.0769](#)
- Smith R. E., et al., 2003, *MNRAS*, **341**, 1311, [arXiv:astro-ph/0207664](#)
- Takahashi R., Sato M., Nishimichi T., Taruya A., Oguri M., 2012, *ApJ*, **761**, 152, [arXiv:1208.2701](#)
- The Dark Energy Survey Collaboration et al., 2015, [arXiv:1507.05552](#)
- Van Daalen M. P., Schaye J., Booth C. M., Dalla Vecchia C., 2011, *MNRAS*, **415**, 3649, [arXiv:1104.1174](#)
- Van Waerbeke L., et al., 2000, *A&A*, **358**, 30, [arXiv:astro-ph/0002500](#)
- Viola M., et al., 2015, *MNRAS*, **452**, 3529, [arXiv:1507.00735](#)
- Vogelsberger M., et al., 2014, *Nature*, **509**, 177, [arXiv:1405.1418](#)
- White M., 2004, *Astroparticle Physics*, **22**, 211, [arXiv:astro-ph/0405593](#)
- Wittman D. M., Tyson J. A., Kirkman D., Dell’Antonio I., Bernstein G., 2000, *Nature*, **405**, 143, [arXiv:astro-ph/0003014](#)
- Zentner A. R., Rudd D. H., Hu W., 2008, *Phys. Rev. D*, **77**, 043507, [arXiv:0709.4029](#)
- Zentner A. R., Semboloni E., Dodelson S., Eifler T., Krause E., Hearin A. P., 2013, *Phys. Rev. D*, **87**, 043509, [arXiv:1212.1177](#)
- Zhan H., Knox L., 2004, *ApJ*, **616**, L75, [arXiv:astro-ph/0409198](#)
- Zuntz J., et al., 2015, *Astronomy and Computing*, **12**, 45, [arXiv:1409.3409](#)

This paper has been typeset from a  $\text{\LaTeX}$  file prepared by the author.